

Use of Biological Standards in Diagnostics Based on mRNA Expression Measurements

Roland Stoughton

Merck/Rosetta

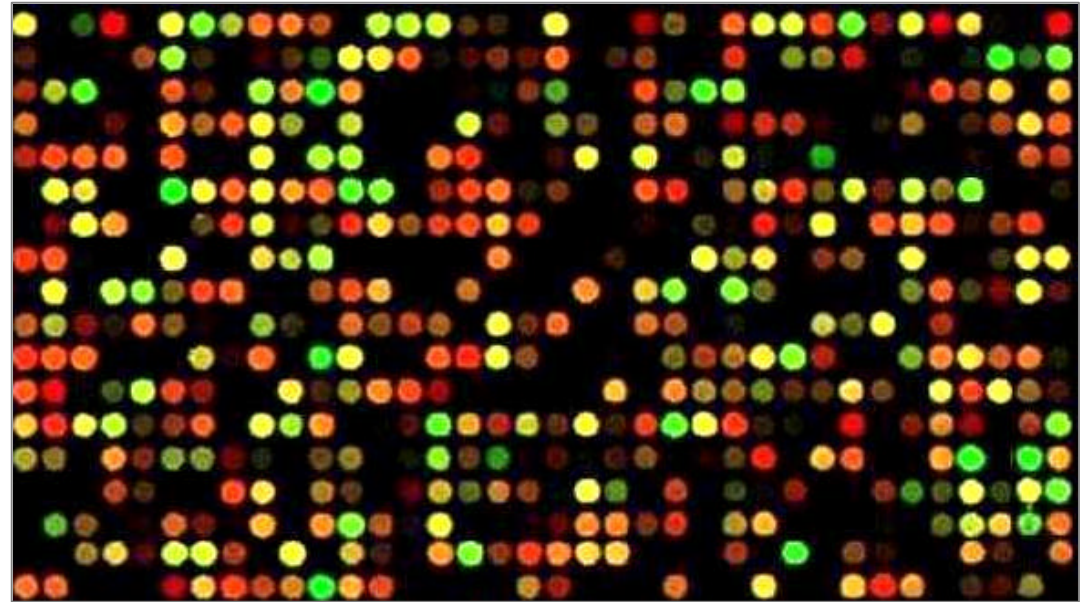
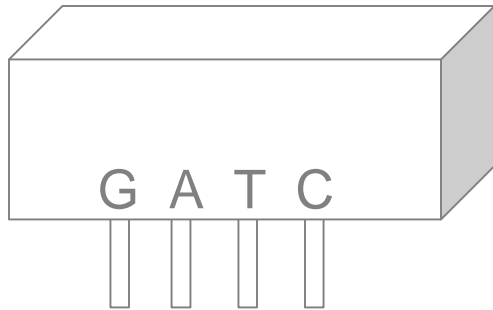
Mar 28, 2003

Use of Biological Standards in Diagnostics Based on mRNA Expression Measurements

- Breast cancer example
- What's really new about these technologies as diagnostics?
- How can standards help?

Breast Cancer Example

Ink-jet *in situ* Oligonucleotide Arrays



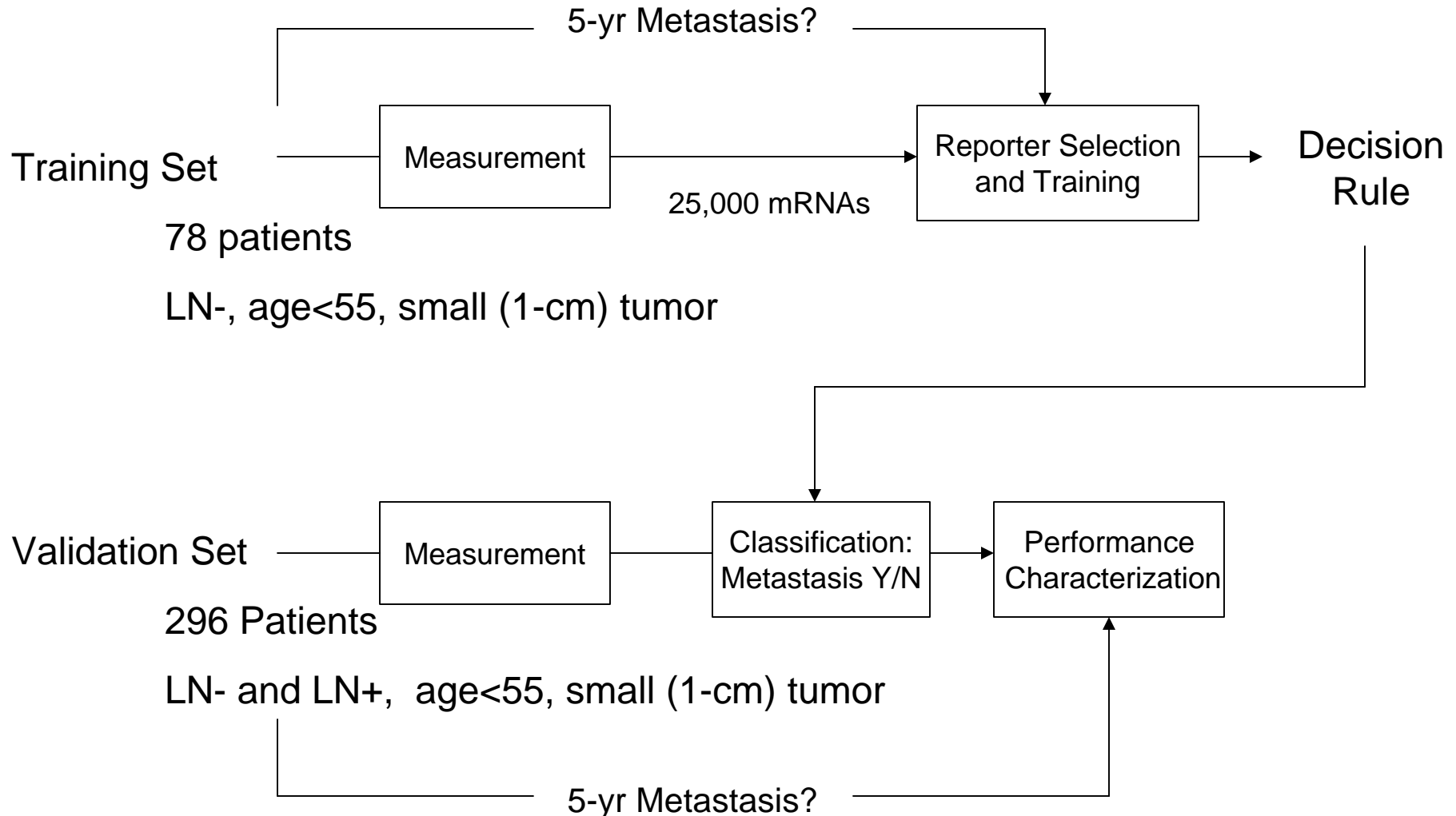
G
A
G
T
C

- 25,000 oligos / 1 x 3 inches
- Any 60-mer at any position
- Two-color hybridizations

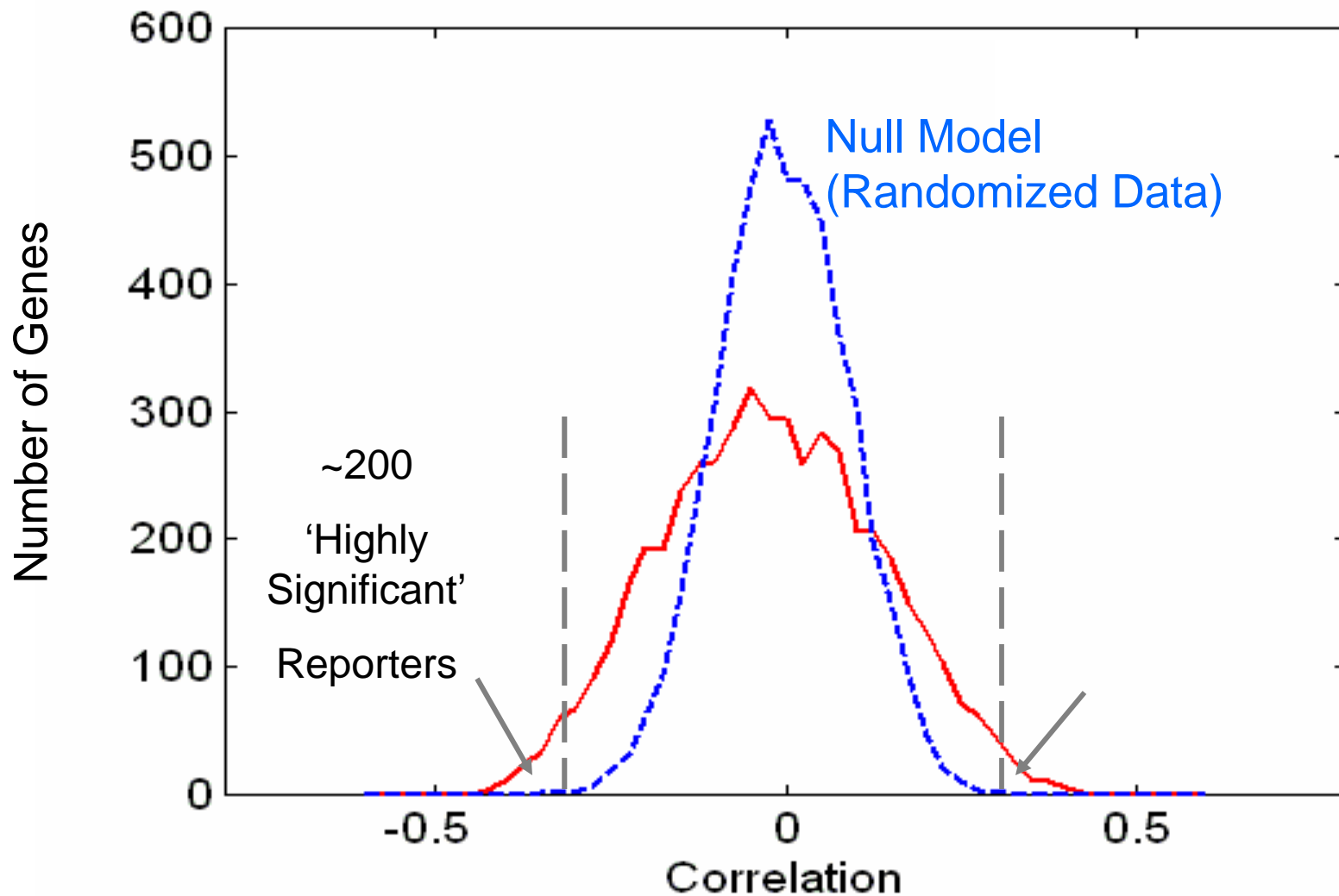
“Red” channel = individual breast tumor

“Green” channel = average pool of all tumors

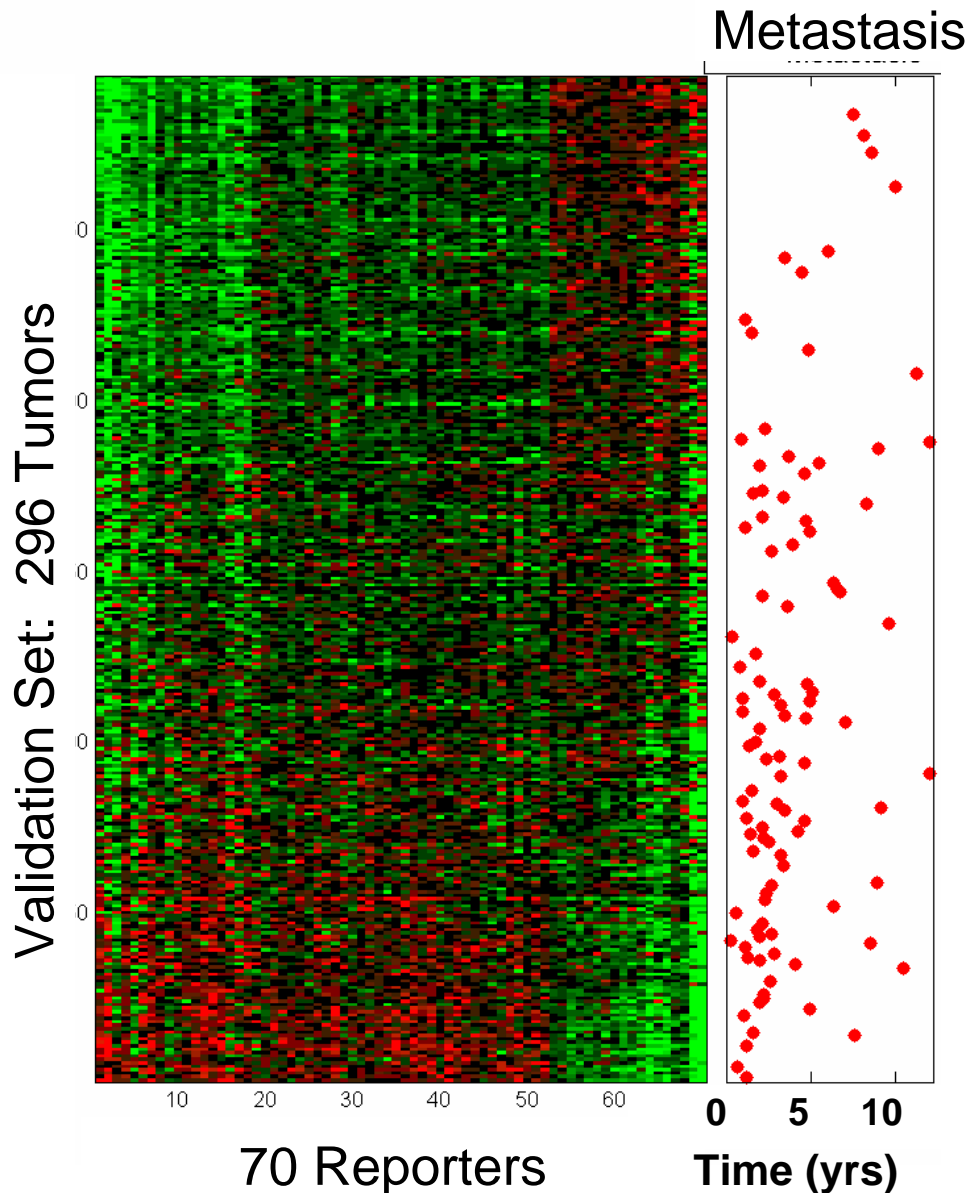
Predicting Breast Tumor Metastasis



Distribution of Correlations of 25,000 mRNAs with the Metastasis Endpoint



Prognostic mRNA Profile for Breast Cancer*

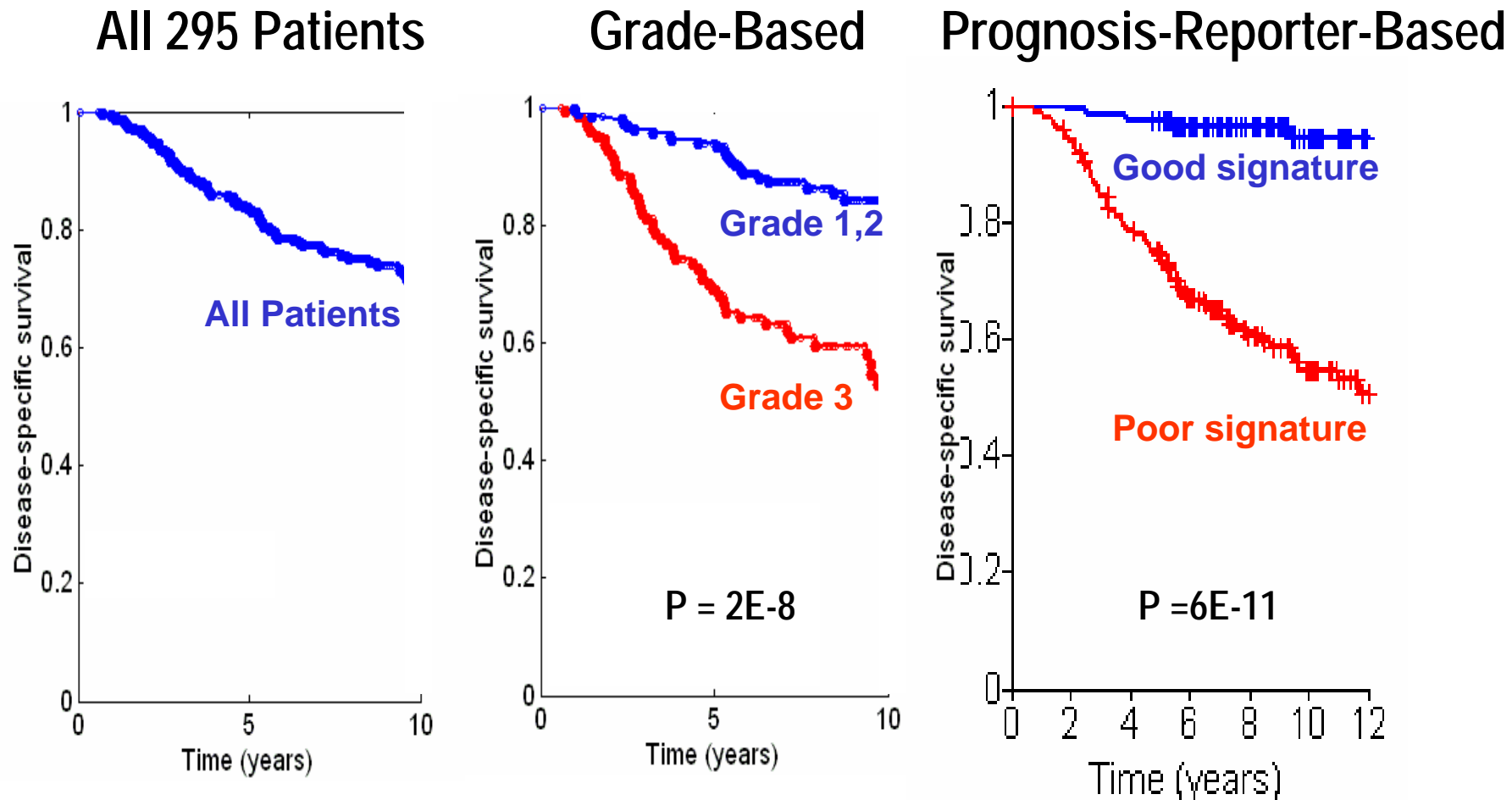


Expression pattern of these 70 reporters is indicative of likelihood and time to metastasis.

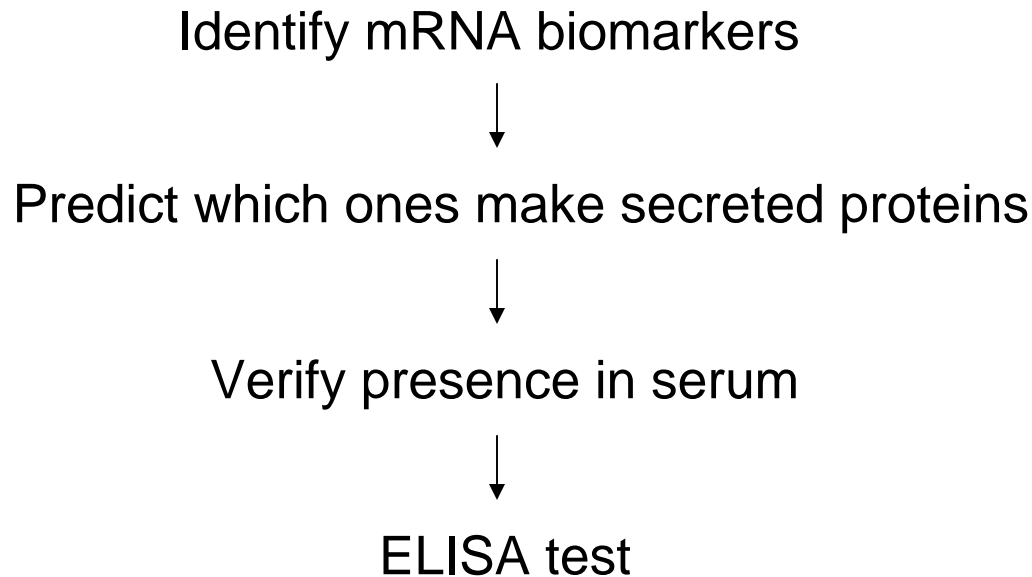
*NEJM, Dec 19, 2002.

Expression profile predictor outperforms existing indicators such as BRCA1 status, tumor grade, etc.

Survival Analysis: Kaplan-Meier Plots



These technologies can help generate diagnostics, even when not used directly as diagnostic platforms

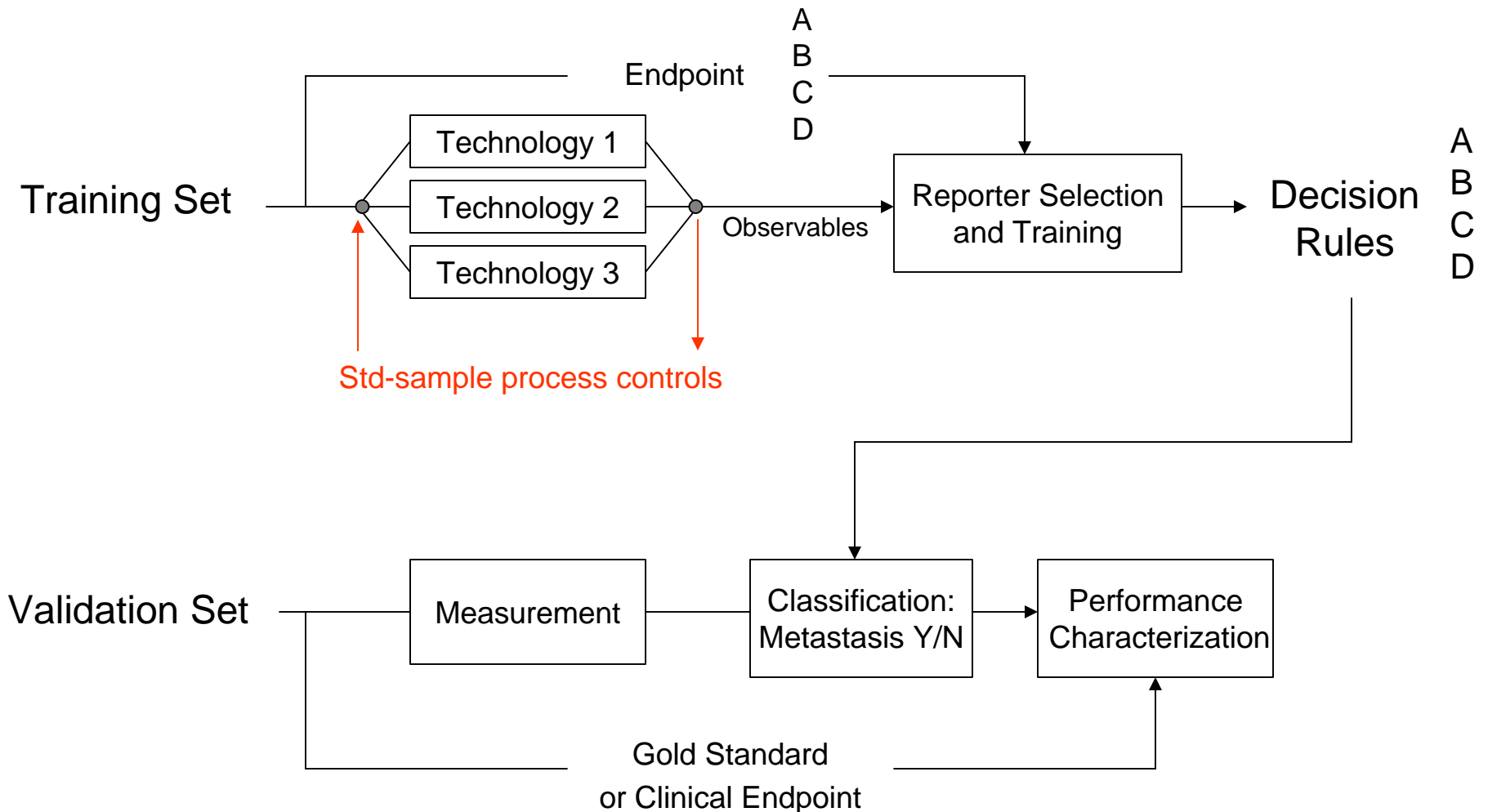


What's really new about these technologies as diagnostics?

What's really new about these technologies as diagnostics?

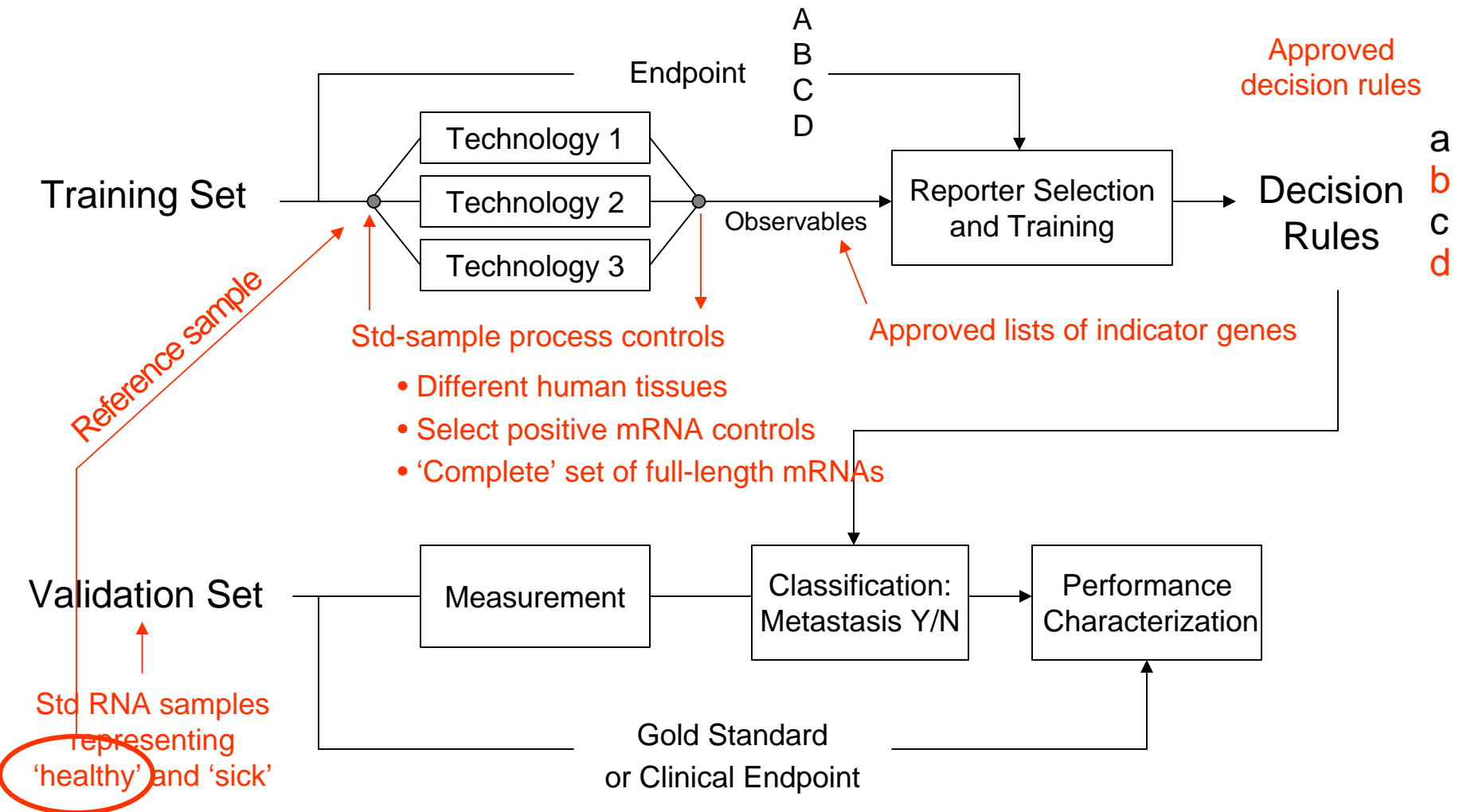
- mRNA
 - Must be separated out from totRNA
 - Susceptible to degradation -- use care and need test
 - Non-circulating (except white cells) -- need the appropriate tissue
 - Each tissue has a different 'normal' abundance distribution -- careful dissection
 - Hybridization tests have good, but finite, specificity -- need to watch out for cross-hybridization
- High-dimensional
 - Many predictive RNAs will *not* have known biological function
 - Decision rules and training algorithms can be complex
 - How to assure robustness
 - How to update as studies accumulate
 - Would like to separate qualification of *measurement* from qualification of *decision rules*

It would be economical to qualify measurements independently of decision rules



How Can Standards Help?

Possible Roles for Standards



'Reference' Sample

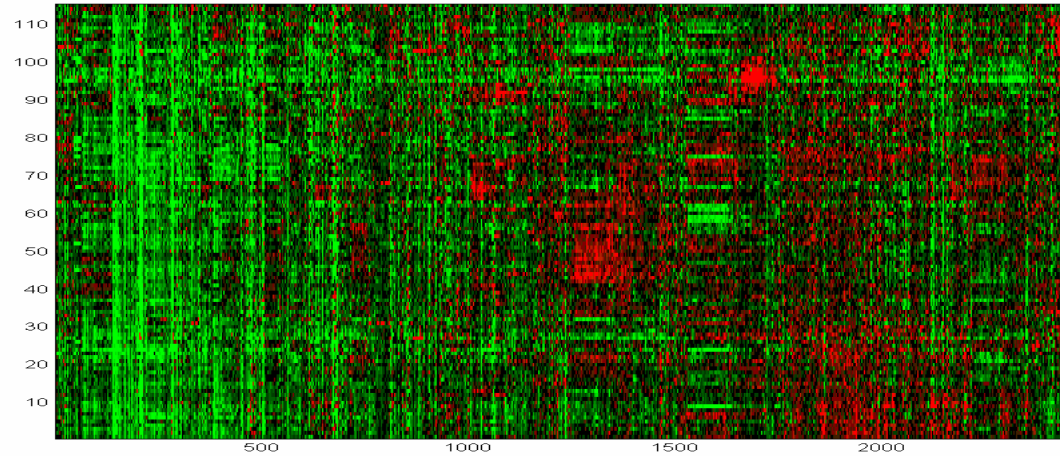
- 'Reference' need not be the 'healthy' state
 - Two-color systems usually need some fixed reference in every hybridization to achieve best accuracy
 - One-color systems may benefit from occasional reference profiles to control for 'drift'
- Although usually not available, an ideal reference would be the *individual patient* in a previous healthy state
 - Controls for genotype, age, environment, ...
 - Reference sample is closely matched -- measurement is more sensitive to subtle changes

Issues with Possible Biological Standards

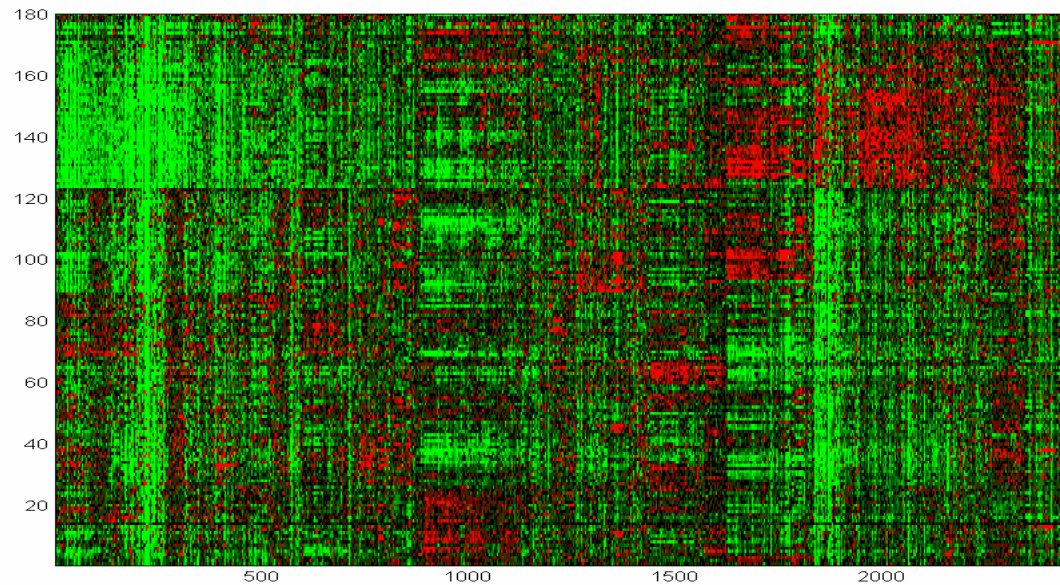
- Control samples representative of each diagnostic endpoint
 - ‘Good’ and ‘Bad’ groups are each heterogenous -- subphenotypes and genotypes
 - Can’t validate a classifier with just a few
- Standard tissue samples
 - What is truth? -- mixtures approach
- mRNA controls
 - Full length mRNAs? Or what part of the molecule? Which splice form?
 - Need ~50,000 clones or more to cover all mRNAs
 - Measure alone, or spiked into realistic sample?

'Good' and 'Poor' Breast Cancer Prognosis Groups are Heterogenous

Good



Poor



Summary

- Canonical development involves
 - measurement technology
 - a training set
 - generation of a classifier
 - a prospective validation set
- Industry will benefit from independent validation of *rules* and of measurement *technologies*
- Standards can be
 - actual RNA samples, or known mRNA mixtures
 - as a biological reference and/or technology platform validation
 - informational
 - approved gene lists
 - approved indicator patterns and decision rules